SAM as PPDG Middleware

Igor Terekhov
FNAL, SAM project, CD, D0

PPDG Collaboration Meeting
13,14 July 2000
Argonne National Lab

Introduction

- History: developed for data access for D0 Run II:
  - ~500TByte data per year;
  - 250 KB/events; 1 GByte files;
  - Total around 1PB
  - 68 institutions; 500 collaborators; 15 (and expanding) countries

SAM Services

- Catalogue all the data files
  - types: 2 dimensions(tier, stream), flexible
  - creator application
  - multiple locations of data files
  - complete access history
- Move the files (retrieve/store) to/from user app., from/to MSS, disks (any valid locations)
  - Access to the underlying MSS
  - Disk caching
- Resource allocation and management

Other SAM Services

- Dataset query manipulation:
  - storing
  - looking up
  - reusing
  - combining
- Straightforward application interfaces
- Error recovery
  - Server state persistent
  - User may restart an activity

Global Data movement through the distributed disk cache

- User app always reads from/writes to local file.
- A user pushes file into (pulls from) SAM and doesn't want to know how/when the system relocates the file.
- Disk allocation at each station en route: where to find room from a distributed pool of disks?
- Every file transfer requires Resource Manager's authorization in order to heed network contention, MSS availability, etc.
- SAM Servers to be introduced below (rather than nodes) form a network. Files must be routed just like packets in traditional network

Distributed Caching: file retrieval in SAM

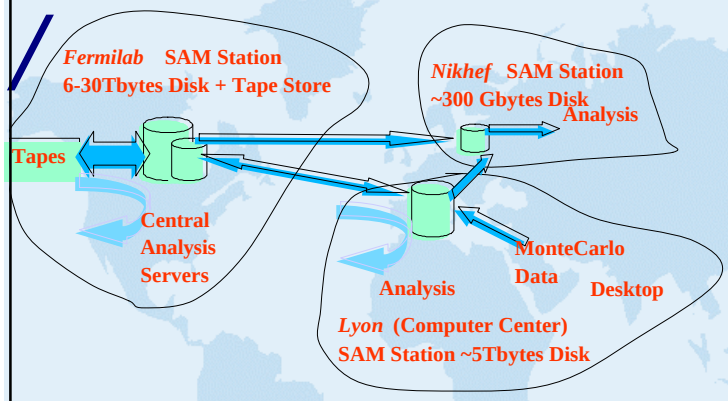
- Station is a collection of resources (CPU, *disk*, etc)
- *Station Master* is a distributed cache manager
- *SM* also runs projects (activity of data retrieval for "consumption")
- Distributed nature of data consumption:
  - a set of machines comprise a station
  - multiple project masters, consumers, processes
  - SM's may get files from each other: **global network**

Distributed Caching: file Storage in SAM

- A station server File Storage Server (FSS) handles requests to import data (raw, processed,...) into SAM
- The system (or physics group administrator) decides file destination
- Intermediate locations may have to be used b/c:
  - "Final" destination is not directly accessible
  - Desirable to keep on "nearby" disks for sooner retrieval
- FSS accepts file storage request, routes the file via other FSS's (stations), **FSS global network**.
- Each interim disk location is managed by station

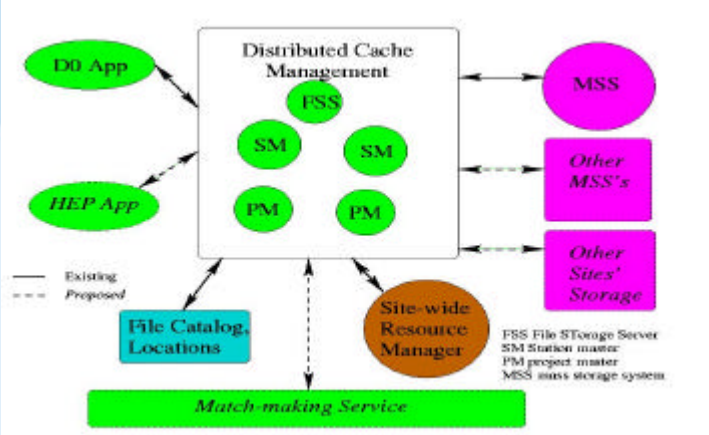
Example of Global data Movement in SAM

- True challenge for a distributed system:
  - import of a file from remote site (INPN Lyon) into SAM MSS
  - the file is on an IDE disk on a desktop
  - the site is connected via *shared* WAN with FNAL
  - MSS at FNAL requires a minimal rate of transfer
  - want to keep it locally (Lyon) as well
  - have friends at CERN who may want the file
  - want to declare it to FNAL central analysis disk
  - handle errors! Any server may restart any time



Distributed Storage and Delivery of Data To Date

SAM distributed caching as a middleware candidate



Summary

- For Grid middleware, SAM offers the capacity of global (fully distributed) disk cache management.
- SAM uses interfaces to external resource managers, data locations "database" and external Mass Storage, possibly Grid components
- SAM provides interfaces to user applications and hides physical data transfer.

More Information

- Ninth Symposium on HDPC http://www.cs.cmu.edu/~hdpc
- SAM home page http://d0db.fnal.gov/sam
- CHEP 2000
  paper http://chep2000.pd.infn.it/paper/pap-c241.pdf
  Presentation http://chep2000.pd.infn.it/pres-c241.ppt